Decision Theoretic Assessment Model for Online Business Games

SANAT BISTA,
CSIRO ICT Centre, Canberra, Australia
and
KESHAV DAHAL AND PETER COWLING
AI Research Centre, University of Bradford Bradford, United Kingdom

Several approaches to reputation and trustworthiness assessment based on probabilistic assessment have considered studying its usefulness in online business environments. It is seen that probabilistic models of reputation and trustworthiness assessment aids in resolving uncertainty by assessing the trustworthiness level of players. In this paper we intend to further enhance the assessment by combining probabilistic assessment with expected utility for each player, thus resulting in a decision theoretic assessment. In this form of assessment, a player makes a decision on the basis of what she believes (given by the probabilistic assessment) and what she wants (given by the utility value of the choice). This method ensures that the players possess a continuous measure of state quality. The assessment of trustworthiness in this model is guided by the principle of maximum expected utility, where a rational player chooses an action only if that action meets her expected utility. Our results show that the decision theoretic models of assessment positively contribute to the evolution of cooperation in a player society. Experiments have been carried out in a business game environment where transactions are modelled using the Iterated Prisoner's Dilemma.

Keywords: Business Games, Decision Theory, Evolution of Cooperation, Iterated Prisoner's Dilemma, Trustworthiness Assessment, Utility Theory

1. INTRODUCTION

In daily life, making a decision on how to choose between particular set of options usually depend on more than one criteria. People tend to compare multiple aspects of each choice and finally reach to a decision by choosing a particular action that seems to maximise their gain. As an example, let us consider a firm which is introducing a new product in the market. Out of many, two parameters that it could assess to reach a decision on whether to introduce the product or not could be: a) the forecast market demand for the product, and b) the current supply of similar products by competitors. A decision to introduce or not should jointly be based on the analysis of both of these parameters (also considering other parameters that apply). It could be unfavourable to reach a decision on the basis of a single parameter only. In the context of online business communities; a buyer's decision on which competing seller to choose for a purchase can be based on the offer price and the seller's reputation. Sometimes, there are situations where the decision-maker must trade-off these sets of parameters. In a more complex form, there can be many parameters (more than two) which need to be considered while making a decision. For example, besides reputation and price, the online buyer's choice of seller can also be influenced by factors like the delivery time, payment methods, specification of the goods, after sales service etc.

A branch of mathematics and statistics that deals with making such decisions is called Decision Theory [Berger 1985]. Decision theory together with Utility theory has been used popularly in both decision making under risk and in decision making under uncertainty [Parsons and Wooldridge 2002; Russell and Norvig 2010]. Risky and uncertain situation represent a partially

Authors' address:Sanat Bista, CSIRO ICT Centre, Canberra-2601, Australia; Keshav Dahal,AI Research Centre, University of Bradford, West Yorkshire BD7 1DP, United Kingdom; Peter Cowling, AI Research Centre, University of Bradford, West Yorkshire BD7 1DP, United Kingdom.

observable task environment where there exists little or no knowledge about the occurrence or non-occurrence of the undesired event or action. The possible action could be inferred through some existing data and probabilistic assessment based on them. The terms risk and uncertainty at times are used synonymously, but the difference lies in the fact that risky situations have some known probability of things going adverse, but with uncertain situation, knowledge of such probabilities are unknown [Riegelsberger, Sasse et al. 2005]. The level of risk involved in a situation increases with the degree of uncertainty existing there. It is thus necessary to tackle uncertainty to reduce the risk and manage the risk to avoid making an undesirable decision. Game Theory [Ross 2009] is a related concept to decision theory which models a decision problem as a competing interaction between the players. A game theoretic representation of a certain decision problem contains a set of players with some defined set of actions and strategies to guide a particular action. Any rational player would choose an action that would maximise the utility value of the game, assuming rational behaviour from the opponent(s).

Our investigation furthers the work to date by adopting a decision theoretic approach to the assessment of a player's trustworthiness in an online business setting. We have considered calculating Maximum Expected Utility (MEU) of any proposed interaction between the players based on expected utility value of the transaction and the probability of cooperation of associated players. This kind of assessment as compared to the one in which we make decisions merely on the basis of past history of cooperation or reputation figure is more realistic and is also scalable to possible future models which might include consideration of additional parameters for evaluation, for example the buyer and seller protection programs in online business. Our results in this paper show that the decision theoretic model favours the evolution of cooperative players.

The remaining part of this paper is organised as follows: Section 2 reviews works in the area of Trust computation, Evolution of Cooperation and Reputation in related application scenarios. Section 3 gives background information on decision theory and utility theory and links our problem with decision theoretic solution techniques. We present our decision theoretic model in section 4. Investigation methods have been outlined in 5 and the results of the experiments detailed in 6. Section 7 gives concluding remarks.

2. RELATED WORKS

Issues related to the evolution of cooperation, Trust and Reputation have been addressed by various authors in [Aberer, Zoran et al. 2006], [Yu and Singh 2000] [Janssen 2006], [Jurca and Flatings 2004]. Aberer in [Aberer, Zoran et al. 2006] outlines the complexity of Trust and Reputation and discusses different approaches to computing trust and reputation. The authors have considered evolutionary approach as one of the many popular approaches that game theorists have been using. In [Yu and Singh 2000] the authors have presented a social mechanism of reputation management in electronic communities. In their discussion around electronic communities, the authors have described the Prisoner's dilemma as a constituent part of the model.

The Prisoner's dilemma has been a very popular framework for investigating behavioural fitness and evolution in a wide range of settings [Axelrod 1987; Frean and Abraham 2001; Yu and Singh 2002; Aberer, Zoran et al. 2006; Janssen 2006; Bista, Dahal et al. 2008]. In one of the most remarkable works on the evolution of cooperation, Axelrod in [Axelrod 1984] addressed fundamental questions about human nature. Starting from basic questions like when should a person be cooperative and when selfish, his work tries to demystify the complexity of cooperation in humans by relating it to the Prisoners' dilemma. Based on Prisoners' dilemma again, Axelrod in [Axelrod 1987] conducted a tournament of strategies using Genetic Algorithms to identify a fittest and an evolutionarily stable strategy.

In a related work [Janssen 2006] has studied the role of reputation scores in the evolution of cooperation in online e-commerce sites. The author discusses whether or not reputation alone can be meaningful in evolving a cooperative society. The paper concludes that high level cooperation is not possible with only reputation scores. In [Chong and Yao 2007] the authors have studied the accuracy of reputation estimation in reflecting strategy behaviours. The authors have illustrated the accuracy to be a function of how previous game memory is used to compute reputation and how frequently the reputation information is updated. [Bista et al 2008], study the role of providing compensation to loss in business games. Their results have shown loss compensation schemes to have positive impact in curbing the defectors but the conclusion is that compensation is not enough a reason to promote cooperation significantly - hence a need for further research in composite models.

The work of identifying trust related parameters and designing suitable trust models for distributed environments have also been of much interest among researchers. The three noted models of trust and reputation in Peer to Peer environment are, EigenTrust [Kamvar, Scholsser et al. 2003], PeerTrust [Xiong and Liu 2004] and PowerTrust [Zhou and Hwang 2007]. [Xiong and Liu 2004] presents a reputation based trust framework relying on transaction feedbacks from peers in a P2P network. The authors have identified trust parameters and defined a general trust metric to define those parameters. EigenTrust [Kamvar, Scholsser et al. 2003] considers developing an algorithm to compute global trust values based on power iterations- thus enabling the system in identifying malicious peers in the network. In [Zhou and Hwang 2007] the authors propose a trust overlay network to model local trust and reveal peer feedback relationships. By selecting a small number of reputable power nodes using a distributed ranking mechanism, the system claims to improve global reputation accuracy and aggregation speed of trust value. Notion of trust is also central to its application in collaborative web applications [West, Chang et al. 2011], Web Services [Dragoni 2010], trust negotiation for semantic web services [Daniel Olmedilla 2004], reputation bootstrapping and management in web services [Malik and Bouguettaya 2009] [Nepal, Malik et al. 2011]. In these domain, trust plays a vital role in predicting web service behaviour and selection [Aljazzaf, Perry et al. 2010].

Though the existing literature has discussed several aspects of trust formulation and reputation based methods for assessment of trustworthiness, they do not particularly discuss computation of trust on the basis of expected utility of each transaction. Our belief is that trustworthiness assessment and a decision as to whether or not to execute a transaction should be based on specific expected utility of the transaction, as a realistic measure for any rational player. Our results show merit in this form of assessment and decision making.

3. DECISION THEORY AND UTILITY THEORY

The decision theoretic framework as a solution to problems involving decision making has been used in many application areas in computer science. Examples of these include multi agent systems MAS [Parsons and Wooldridge 2002], page ordering and ranking systems [Cohen, Schapire et al. 1998; Zoeter, Taylor et al. 2008], multimedia content identification [Avinash, Ashwin et al. 2008], AI planning [Blythe 1999], Information retrieval [Ye Diana and Guisseppi 2006]. The potential of the decision theoretic framework in modelling problems related to representation, inference, and knowledge engineering was described by [Horvitz, Breese et al. 1988] in the context of expert systems, and today most of the popular applications are in these domain.

Decision theory basically relies upon probability theory and utility theory, where preferences expressed by utilities are combined with probabilities to reach a decision. This simplifies decision making to a simple act of choosing the alternative on the basis of the utility score assigned to different outcomes. The fundamental idea of decision theory is that a decision is rational if and only if it has chosen an action that yields the highest expected utility averaged over all the possible outcomes of the action [Russell and Norvig 1998]. This is referred to as the principle of Maximum Expected Utility (MEU). The concept of utility theory can be used in cooperative as well as non-cooperative games [Kaneko and Wooders 2004].

Utility theory and in particular the use of expected utility maximization has been a predominant analysis method for decision making under risk [Kahneman and Tversky 1979]. Accepted as a normative model of rational choice the theory is widely applied as a descriptive model of economic behaviour [Friedman and Savage 1948; Schoemaker 1982]. In utility theory, the choice of an action is based upon a real valued score which is produced by utility functions that represent the desirability of the particular state. Expected utility theory could also be thought of as a "probability-weighted utility theory" where each alternative is assigned a weighted average of its utility values under different states of nature using probabilities of these states as weights [Hansson 1994].

According to [Russell and Norvig 1998], if U(S) is the expected utility and P(S) represents the probability of the occurrence of any state S, then the maximum expected utility for all the n possible states of S is calculated as :

$$\bigcup([p_1, S_1; \ldots ; p_n, S_n]) = \sum_i P(S_i)U(S_i) \tag{1}$$

Let us take a simple example here to illustrate the concept. Suppose that an individual investor with a current wealth of £10 is considering an investment. A fair coin (which has equal probability of showing head (0.5) or tail (0.5)) is flipped in this investment. If head comes up then the investor wins £5, and his total wealth increases to £15. In case of a tail, the investor looses £5 and his wealth decrease to £5. Thus, the two states possible in this case are first the head state which represents a gain and the tail one which represents loss.

The maximum expected utility (MEU) of this investment according to equation 1 is be obtained as:

MEU=0.5 X (£5)+0.5 X (- £5)=£0

This example is a fair game [Weisstein 2009]. In real world problems, where there exists some knowledge on the probability of occurrence of some event, the situation can be different. In such cases, the probability of occurrence, which is 0.5 on either side of a fair game, will be replaced by respective probabilities. An example of such situation is the transactions over eBay, where a buyer can have information on the reliability of a seller through the seller's feedback profile.

When modelling business transactions, it is natural to evaluate benefits in terms of monetary value. In our models defines in the subsequent sections, we use money as an attribute while measuring utility. In such cases maximum expected utility is also referred to as Expected Monetary Value (EMV) [Jordaan 2005]. Research has shown that utility of money is proportional to the natural logarithm of the amount of money [Kritzman 1998; Jordaan 2005; Russell and Norvig 2010]. Figure 1 presents a plot of monetary value (1 to 50) against its natural logarithm to produce the utility curve.

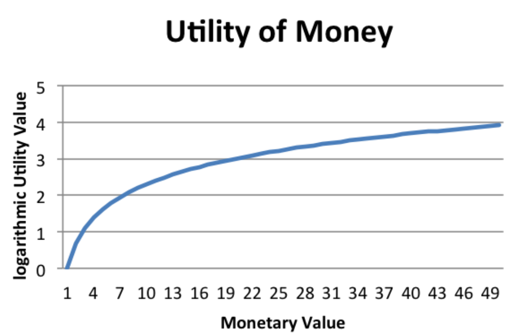

Figure 1: Utility Curve of Monetary Value (1-50).

The curve of monetary utility fulfils both the properties of utility function namely the non-satiation property and the risk aversion property - thus making it a legitimate utility function. The utility function of money is a twice differentiable function (Bianconi 2003). For any monetary value V>0, the first derivate of the utility function U is always greater than zero (i.e. U'(v)>0). The second derivative of the function is always less than zero (U"(v)<0). The non-satiation property observed by the first derivative suggests that the utility increases with increasing monetary value. The risk aversion property observed by the second derivative suggests that the marginal utility decreases with the increasing wealth. This means that for any rational player, there lies always a desire to obtain a little more wealth, but with accumulating wealth, the usefulness of the unit value of wealth decreases [Russell and Norvig 1998; Bianconi 2003].

In the next section, we discuss our decision theoretic model and the trustworthiness assessment model which is applied to a business game scenario based on the Iterated Prisoner's Dilemma model.

4. DECISION THEORETIC MODEL

There are two particular instances where we use decision theoretic principles in our business game model. Firstly, we design the payoff of each interaction to be based on the utility value measured in terms of units called utils. The second is in calculating the trustworthiness threshold. Rather than reputation scores of players determining whether or not the game should be played, we base our decision to do so on the maximum expected utility (MEU) value - explained in section 3 of this paper. The concept and design of the utility payoff model is explained in the following subsection.

4.1 The Utility Payoff Model

The utility pay-off is concerned with the outcome state of each interaction, and depends on the action taken by particular players in the game. The model we use here represents the business game between the players using the principles of the Iterated Prisoner's Dilemma (IPD) model. The IPD (and also our business game model) represents a two-player, non-zero sum, non-cooperative, simultaneous game with players following pure strategy for interaction [Chaudhuri, Sopher et al. 2002; Ba, Whintson et al. 2003; Li 2004; Howley and Riordan 2007].

Our game theoretic model consists of two players, Player One (P1) and Player Two (P2). At each step of the interaction between players, each player has two possible actions: Cooperate or Defect.

Four different situations can arise: first, both players cooperate, second, when the first cooperates and the second defects, third, when the second cooperates and the first defects, and finally when both defect. Each of these actions has an associated pay-off. Pay-offs can be assigned arbitrarily, or they can be based on the gain and losses according to the nature of the game. For example the pay-off matrix presented in Table I lists the possible outcomes for a typical business game where Player One is a Seller and Player Two a Buyer. The symbol γ represents the price of the goods in transaction.

Table I. Payoff Matrix for a Typical Business Game

		Buyer	
		Cooperate (C)	Defect (D)
Seller	C	Rseller= γ Rbuyer= γ	Sseller = - γ Tbuyer = 2γ
	D	Tseller = 2γ Sbuyer = - γ	Pseller = 0 Pbuyer = 0

The values for the pay off are typically called Reward for cooperation (R), Temptation to defect (T), Sucker's Payoff (S) for cooperating against a defecting player, and Punishment for mutual defection (P). For the dilemma to hold, the following condition must hold true (Axelrod 1984):

$$T>R>P>S \tag{2}$$

This inequality ensures that cooperation is pareto optimum and that defection is the equilibrium play [Axelrod 1984]. In the example above, these values correspond to the benefit that each would derive from the respective actions. There is further an interesting consequence if the game is being played for more than one round. Typically known as Iterated Prisoner's Dilemma (IPD), in such games the collective pay off obtained through continuous cooperation can be greater than that obtained through deceptive acts. Thus an additional condition given below must also hold true in this case [Axelrod 1984] [Kuhn 2007].

$$2R>T+S \tag{3}$$

Equation (3) ensures that full cooperation is optimal and can outclass swinging cooperation-defection actions in repeated games [Dawkins 1989]. The situation in the Iterated Prisoner's Dilemma is also a useful model in an online business environment [Yu and Singh 2000; Aberer, Zoran et al. 2006]. If we consider an online trading environment where completely anonymous buyers and sellers interact, they share a dilemma as to whether or not they should reciprocally cooperate. If the seller sends the good and the buyer sends money worth the goods, both of them are rewarded and the transaction meets a happy ending. If either of them defects while the other cooperates, then the cooperator is badly hurt in the transaction and the defector scores highest (at least in the short term). If both defect then each scores a payoff better than sucker's score (and interestingly equal to the reward in value as both retain the money and goods with them), but is nevertheless going to be a bad reference for future transaction. Further, a payoff for a both defect scenario cannot be considered equal to Reward, as there has been no transaction at all resulting in a status quo. A zero return for no transaction is what the payoff in this case would be. [Ba et al. 2003] have considered a zero return in such scenario.

In our work, we consider the Prisoners' dilemma as a model for pay-off in a business interaction between two decision theoretic players. We use the term player to mean a seller or buyer participating in the business game, and further assume that the goods and equivalent wealth have the same level of significance to both of the players (seller and buyer). In game theoretic modelling, representation of gain and loss in business and investment scenario through pay off based on goods price and investment amount have been practiced in [Chaudhuri, Sopher et al. 2002; Ba, Whintson et al. 2003]. We base our pay-offs on the cost price of the goods, but carry it further on to calculate gain and loss in terms of the maximum expected utility(MEU) of the player. As described in section 3, the maximum expected utility of a player is a function of the expected utility of the trade (the cost price of the goods in this case), and the expected probability of cooperation of the opponent player at that particular instance. The consideration of MEU as a payoff as opposed to the exact monetary value of the goods is more realistic and is the heart of the decision theoretic model that we incorporate in the game. The probability based payoff in our case introduces dynamic value to the gain and loss, thus truly preserving the subjective nature of the utility of transactions. Later, we show through our results that this approach has merit as it contributes in increasing the overall cooperativeness of the population, and also in optimising the average utility of business games.

To explain the model, let

the value γ represent the price of the goods in transaction

$p_{1,t}$ represent the probability of cooperation of player one (P1) at time t

$p_{2,t}$ represent the probability of cooperation of player two (P2) at time t

M_{P1} represent the maximum expected utility of the transaction for P1

M_{P2} represent the maximum expected utility of the transaction for P2

$R_{P1,t}$ represent the reward pay-off to P1 for mutual cooperation at time t

$R_{P2,t}$ represent the reward pay-off to P2 for mutual cooperation at time t

$T_{P1,t}$ represent the temptation payoff that P1 obtains while defecting against the cooperating P2 at time t

$T_{P2,t}$ represent the temptation payoff that P2 obtains while defecting against the cooperating P1 at time t

$S_{P1,t}$ represent the sucker's payoff that P1 obtains for cooperating against a defecting P2 at time t

$S_{P2,t}$ represent the sucker's payoff that P2 obtains for cooperating a defaulting P1 at time t

$P_{P1,t}$ represent the punishment pay-off for mutual defection for P1 at time t

$P_{P2,t}$ represent the punishment pay-off for mutual defection for P2 at time t

On the basis of this, the maximum expected utility of a transaction (M) for any player P (with two possible outcome states of cooperation, or defection) according to equation 1 can be obtained as:

$$M = p * \gamma + (1 - p) * (-\gamma) = (2p - 1)\gamma \tag{4}$$

Further, the reward, sucker, temptation and punishment payoff for these players on the basis of maximum expected utility relation in equation 4 can be obtained as follows:

$$R_{P1,t} = (2p_{2,t} - 1)\gamma \tag{5}$$

$$R_{P2,t} = (2p_{1,t} - 1)\gamma \tag{6}$$

$$T_{P1,t} = \gamma + (2p_{2,t} - 1)\gamma \tag{7}$$

$$T_{P2,t} = \gamma + (2p_{1,t} - 1)\gamma \tag{8}$$

$$S_{P1,t} = -((2p_{2,t} - 1)\gamma) \tag{9}$$

$$S_{P2,t} = -((2p_{1,t} - 1)\gamma) \tag{10}$$

$$P_{P1,t} = P_{P2,t} = 0 \tag{11}$$

The models regard time as significant. This is mainly to represent the dynamic nature of each of these models. A utility based on cost alone remains static throughout the game, but when the payoff becomes a function of the cost and cooperation probability, due to the fact that reputation of players is dynamic, the whole computation leads to different values for each interaction. The pseudo code in Figure 2 presents an algorithm based on which individual players are awarded payoffs for a transaction.

The pay off for the punishment case is kept at a zero value to represent the status quo, this was reasoned earlier in the same subsection. The payoffs honour the inequalities (presented in equations 2 and 3) necessary for a dilemma to exist.

4.2 The Utility Threshold Model

Another important aspect of a decision theoretic design is the threshold model. By threshold, we mean the minimum expected utility that is believed to be obtained through certain transaction. In a reputation model, this threshold is equivalent to the minimum expected trustworthiness of the player in order to play the game. A decision theoretic threshold assessment couples the

```
Function DT-awardPayoff (actions) returns payoff

Define: Payoff : M_P1, M_P2

        Moves: Cooperate(C), Defect(D)

IF P1.move=C & P2.move=C THEN

        M_P1 ← R_P1,t , M_P2 ← R_P2,t

IF P1.move=C & P2.move=D THEN

        M_P1 ← S_P1,t , M_P2 ← T_P2,t

IF P1.move=D & P2.move=C THEN

        M_P1 ← T_P1,t , M_P2 ← S_P2,t

IF P1.move=D & P2.move=D THEN

        M_P1 ← P_P1,t , M_P2 ← P_P2,t

return payoff;
```

Figure 2: Pseudo code for the Payoff Model

probabilistic assessment with the utility value of each transaction to set a desired expectation level in order to play a game. Essentially, a threshold in this case is a filter that allows or disallows an engagement based on whether the opponent meets or doesn't meet the minimum expected utility.

Let T represent the threshold level for a transaction, and U (λ) represent the utility of any product worth λ. Given p is the probability that a transaction goes successful, a maximum expected utility based threshold T is given by:

$$T = p * U(\lambda) + (1 - p) * (-U(\lambda)) \tag{12}$$

Just before the transaction, each of the players calculates the maximum gain that could be obtained through the transaction by using the formula given in equation 4. Each then compares this gain M with the threshold barrier T obtained through equation 12. A transaction happens only when the gain is greater than or equal to the T value. The pseudo code in Figure 3 presents this algorithm.

```
Function playGame (Threshold) returns Boolean

Define: Threshold : T

        Maximum Expected Utility of P1: M_P1

        Maximum Expected Utility of P2: M_P2

Compute (M_P1, M_P2);

Compare (M_P1, T);

Compare (M_P2, T);

IF (M_P1 ≥ T) and (M_P2 ≥ T)

        Return (True: PlayGame)

Else

        Return (False: noGame);
```

Figure 3:Pseudo code for Threshold Computation and Game Decision.

The method of investigation outlined in section 5 makes use of these two models designed in this section. The different parameters and values for them are also discussed and explained in the following section.

5. METHOD OF INVESTIGATION

In section 4.1 we explained a prisoner's dilemma based business game model. To simulate this environment, we use a Spatial Iterated Prisoner's Dilemma (SIPD) environment. In spatial IPD the players are arranged in some "geographical" positions and most of the interactions take place between the neighbouring players [Kuhn 2007]. Such arrangements exhibit games in clusters thus providing a natural representation of our problem [Masuda and Aihara 2003]. As an example, in an eBay-like setting, sales items are put under several categories and each category has some defined sets of sellers who interact with buyers sharing similar interest. This is in contrast to tournament-like fitness selection approaches where interactions happen between every possible player. Our simulation model consists of a two dimensional spatial grid of players on a toroidal surface. Each player has eight neighbours with whom it plays the cooperation-defection game. This represents a selective game playing environment where each player doesn't play with every player, but instead the games happen in cluster- thus providing a natural representation to the business game problems involving different cluster of buyer and seller. Figure 4 shows a sample 5x5 square grid with 25 players in it. The highlighted player P7 plays games for a specified number of times with its eight neighbours P1, P2, P3, P8, P13, P12, P11, and P6.

P0	P1	P2	P3	P4
P5	P6	P7	P8	P9
P10	P11	P12	P13	P14
P15	P16	P17	P18	P19
P20	P21	P22	P23	P24

Figure 4: A 5x5 spatial grid with player P7 highlighting its 8 neighbours

Players are characterised by strategies for a 3-memory game. This means that each player while interacting with its neighbour will keep track of past three games. An action on what to do next will be guided by the particular strategy that matches the past three histories of interaction that each player maintains. Our choice of 3-memory has to do with strategy representation. We need the representation to be comprehensive such that it is capable of representing most of the standard strategies such as Tit-for-Tat, all time cooperator, all time defector, Tit-for-two tats, Pavlov etc and more of the random strategies.

The evolution in the game which we study uses a Genetic Algorithm. Genetic Algorithms (GAs) as an intelligent search technique were introduced by John Holland [Holland 1975; Axelrod 1987; Goldberg 1989]. Our use of GA evolution is based on, and inspired by, the original work of Axelrod in [Axelrod 1987]. In the GA environment, a player strategy is represented by a chromosome as a fixed length representation of the possible actions for a three memory game. With two possible actions (Cooperate or Defect) that can be played by each of the two interacting players, the possible moves are CC, CD, DC, and DD. A three memory strategy for these four possible moves needs a 4x4x4=64 bit chromosome length. Beside Axelrod, Jenifer Goldbeck in [Goldbeck 2002] has also used similar encodings to study the evolution of strategies in Prisoner's Dilemma. [Axelrod 1987] used additional 6 bits as an assumption to determine the first three moves. A variation of this was used by [Errity 2003] and we are following the same scheme of additional bit encoding, in which 7 extra bits are used for encoding actions for the first three relative moves (relative to opponent moves). It is thus not required to encode an assumption of the pre-game history. This makes the total chromosome length to be 71 bits with each locus outlining an action C or D to perform, depending on the last 3 game histories.

The model when initialized consists of a specified number of players with random strategies. In each generation of evolution a pre-specified number of games are played between two players, at the end of which, the fitness of the chromosome of each player is evaluated by referring to the payoff values for the game. Within the set of eight interacting players the first player is selected using roulette wheel. The second player chosen for mating would be the fittest in the neighbourhood. One-point crossover [Poli and Langdon 1997] is used to break the parent chromosomes at the same random point. Linear scaling of the fitness as described in [Goldberg 1989] has been used to prevent premature convergence of the evolution.

Cooperativeness of an individual strategy is assessed through its overall probability of cooperation, which in turn relies upon the number of references in the chromosome guiding a cooperative action. Strategies are classified into six different types as per the values of cooperation probability listed in table II. Our classification at the moment is arbitrary (for the purpose of testing the model); however this can be easily mapped to any existing real world benchmarks when necessary. For instance, a seller with even 90% of positive feedback score in eBay might be considered untrustworthy for transactions. By suitably studying the characteristics of the application environment, thresholds for player types can be defined. The classification we are considering here is sufficient and descriptive enough to study the evolution trend of cooperativeness in our experiments.

Table II. Classification of the Strategy based on cooperation probability

Strategy Type	Cooperation Range in Chromosome	Probability of Cooperation
Very Cooperative	> 65	>0.91
Cooperative	55 to 65	0.77 to 0.90
Good	50 to 54	0.7 to 0.76
Okay	45 to 49	0.63 to 0.69
Dishonest	35 to 44	0.49 to 0.62
Very Dishonest	< 35	< 0.49

In addition the following parameters in Table III (and values related to evolution) were used in the experiments:

Table III. Simulation Parameters

Parameter	Value
Total number of Players	10 x10 matrix =100
Generations of Evolution	100
Transactions per Generation	151
Crossover Probability	0.98
Mutation Probability	0.01

The player population of 100 was chosen to represent a good enough size of strategies. With this number of players, sufficient interactions were observed, as each player in the grid plays with eight neighbours and each of these neighbours in turn play with their eight neighbours. Simulations were limited to 100 generations as at this point, the evolution was relatively stable in terms of the average pay-off derived from each interaction. We had to be careful with the number of transactions as this number was proportional to the time taken by the simulation. However, a small value for transaction frequency might prevent the players from exploiting most of their memory three strategies. Axelrod in [Axelrod 1987] had used an average run length of 151 for his strategy tournament. We took the same value of 151 as a tradeoff between performance and fitness. An important aspect of the simulation is the initial composition of players. In our experiments we have chosen a balanced population of strategies where a larger number of players would act in a balanced way and a smaller number of them would act as either of the extreme ones. Our composition of the players is not based on any particular application scenario, but it is rational to assume that many players in society would play balanced games and cheaters as well as blind co-operators would remain a minority.

When the game starts between the players, pay offs are derived based on the utility pay off model described in 4.1 and a decision on whether to play a game or not is guided by the utility threshold model in section 4.2. In our experiments, we assume all the goods transacted to be of the same value, worth £10. We make this factor a constant so as to produce comparative results. To understand how maximum expected utility (MEU) grows against the increasing probability of cooperation from 0 to 1, we compute MEU for a utility value of 10. The computation is made using equation 1. The examples presented in section 3 had shown how to compute MEU for any given utility value. The chart presented in Figure 5 shows this distribution.

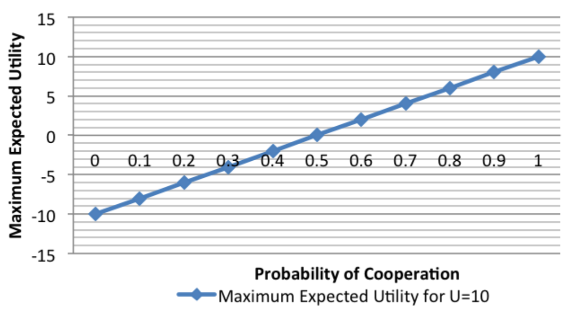

Figure 5: Growth of MEU against increasing cooperation probability for a utility value of 10.

From the figure, we can see that for a single valued product, its utility increases with the increasing probability of cooperation.

We have considered three different threshold levels for the investigation; first, a threshold of zero representing a fair game expectation. Secondly, a mildly strict threshold with an expected utility of 5 or higher, and third a strict threshold demanding the players to cross the barrier of 9 (values calculated with the cost being set at £10). In the experiment, the probability of cooperation that is used in calculating the maximum expected utility is dynamic and changes with the behaviour of particular player in the game. At the beginning, players develop some game history by first warming-up, and during this time the threshold computation model doesn't come into effect. This warm up is required to let the players build some reputation so that they meet the basic threshold requirement after the decision model comes into effect.

6. RESULTS

To test the effectiveness of the decision theoretic model in promoting the cooperative evolution in the player society, we have selected a set of metrics which are observed to study their behaviour as evolution progresses. The sub-section below describes them.

6.1 The Observation Metrics

i. Evolution of Cooperative Strategies: Table II in section 5 presented the classification of the players according to the cooperation probability of their strategies. Out of the six different types, the first three (very cooperative, cooperative and good) of them have been classified as cooperative ones. Thus, a cooperative player in this context would have a cooperation probability of 0.7 or higher. We study the evolution of the proportion of cooperative player strategies and see whether the new setting has been favourable for their growth.

ii. Evolution of Cooperativeness Index: The Cooperativeness Index is a ratio of the expected probability of cooperation of the strongest strategy to that of the weakest strategy in each generation. Any strategy that could obtain a highest pay-off was termed strongest while the one leading to lowest pay-off was termed weakest.

For an ith strongest strategy Si and jth weakest strategy W_j, the cooperativeness index (CoOPIndex) is given by:

$$CoOPIndex = \frac{p(S_i)}{p(W_j)} \tag{13}$$

Where, p represents the expected probability of cooperation.

The measure of cooperativeness index gives an indication of how cooperative the strongest strategy is. By default, defection earns the highest pay-off, so there is a likelihood that any strongest strategy consists of a majority of defective actions, but this is exactly what we don't want to see in a cooperative society. We are trying to see here that the strongest strategy is also a cooperative strategy.

iii. Evolution of average utility payoff: Each interaction in the game has an associated payoff. In each generation of evolution, the games are played for a specified number of times before the fitness of an individual strategy is evaluated. Any strategy earning the highest payoff is considered the fittest for reproduction. In this context, an average utility payoff represents the average of the payoffs of the strongest and weakest strategy in each generation. Evolution of this figure relates itself to the evolution of cooperation as we see this in the results discussed below.

6.2 Results and Analysis

Firstly we consider the evolution of cooperative strategies. As mentioned before, strategies with the probability of cooperation with 0.7 or higher were considered at the beginning. With the interesting trend in the evolution, we further restricted it to see the evolution trend of those with a probability of 0.77 or higher (including just very cooperative and cooperative players). Figure 6 and 7 present this evolution.

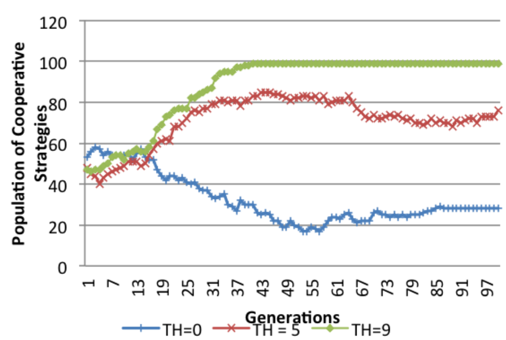

Figure 6: Evolution of cooperative players in a decision theoretic setting at different threshold levels.

From Figure 6 it can be seen that at the maximum expected utility threshold of 9, the evolution of cooperative players soon reached the full size of 100% before 50th generation itself. This shows the effectiveness and usefulness of the trustworthiness threshold in promoting cooperation in the society. With no threshold (TH=0), evolution of cooperative players is absent. At the end of 100th generation, the cooperative population is seen to be almost negligible. At threshold of TH=5, the evolution is better than at zero, but it still remained below 80%.

As the cooperative population reached a 100%, we slightly increased the level of cooperativeness of the strategies to 0.77 in order to see whether the evolution trend for even better cooperators did match with that in the earlier case. The trend lines in Figure 7 show that they are in line to the trend seen in Figure 6.

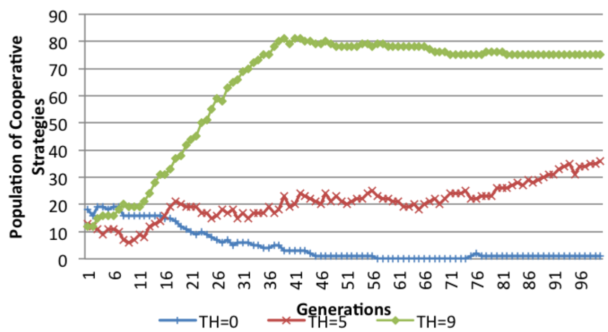

Figure 7: Evolution of cooperative players with cooperation probability ≥ 0.77 in a decision theoretic setting at different Threshold levels

Thus, the result here suggests that the decision theoretic model applied to the simulation system have played a constructive role in promoting cooperation. The increase of the cooperative population with the increase in strictness of the threshold value shows efficiency of the system.

The second thing that we asses to evaluate the systems efficiency is the cooperativeness index (CoopIndex). Population study limits itself to the study of certain efficiency only. With cooperativeness index, contribution of the cooperativeness of weakest strategy is also considered. This means that cooperativeness index and the trend in its evolution are much expressive forms of the evolution of cooperation. Figures 8, 9 and 10 show the evolution of cooperativeness at the three different threshold levels. The trend lines have been presented separately for better visuals.

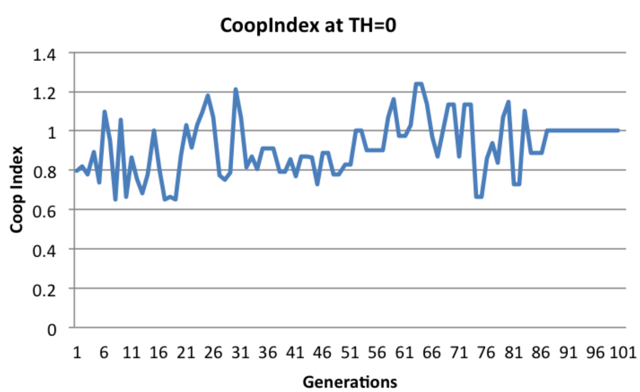

Figure 8: Evolution of CoopIndex at TH =0

The first trend line in Figure 8 shows a relatively fluctuating line as compared to the other two thresholds. In terms of the levels of fluctuations, they are seen to be decreasing with the increasing

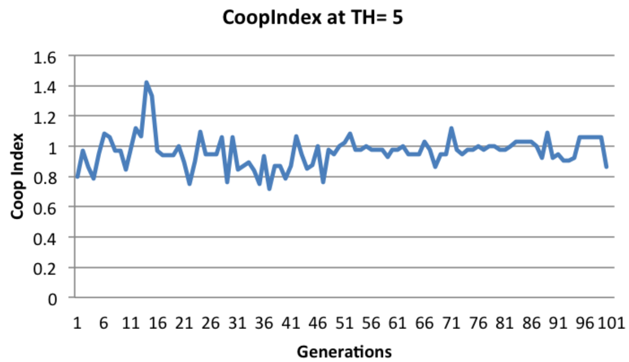

Figure 9: Evolution of CoopIndex at TH =5.

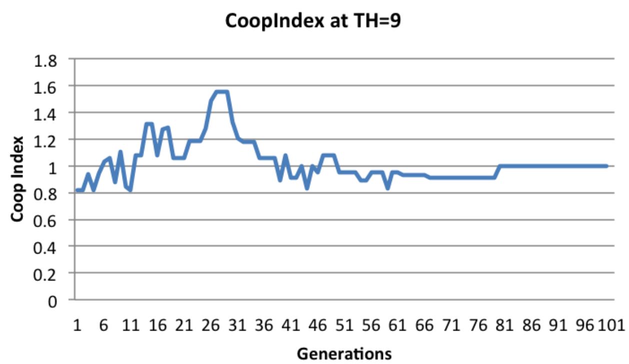

Figure 10: Evolution of CoopIndex at TH =9

threshold levels. Fluctuations represent inconsistencies in the evolving strategies, hence the lesser the fluctuation the better the results.

Besides fluctuation, another thing to look for is whether or not index lies at or above the value of one. Values lower than one signifies that cooperating instances in the weakest strategy are greater than that in the strongest strategy. The results show that in case of a higher expectation of TH=9, the index has maintained a value greater than one most of the times, and that it has assumed a more stable value of one towards the later phase of evolution. The average of the indexes through generations presented in Figure 11 shows that at TH=9 the average itself is above one.

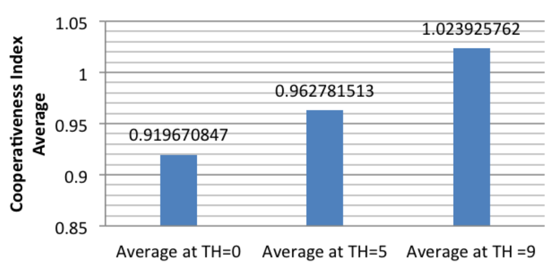

Figure 11: The average CoopIndex at different threshold levels

Finally, we analyse the average utility evolution. Average payoffs directly represent the gain through each transaction, but indirectly they can also be interpreted to predict what actions among the players might have been dominant. A very high average payoff, which is greater than the one that would be obtained through mutual cooperation suggests defection for cooperation to be a dominant strategy- as this gives a very high temptation payoff. At the same time very low average values can be interpreted as saying that defection for defection might have been the prevailing actions among the players.

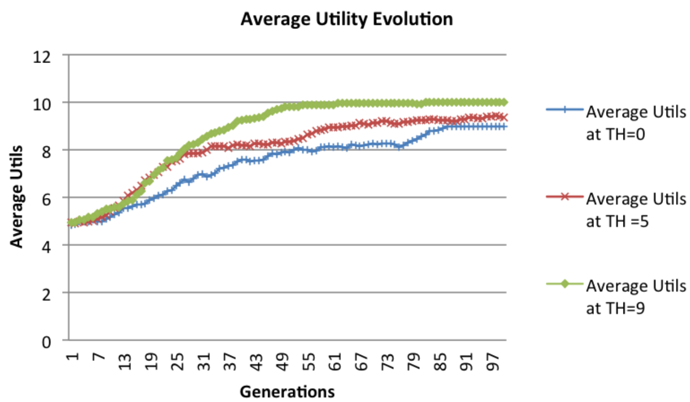

Figure 12: Evolution of average utility at different thresholds.

The trend in the average utility evolution in Figure 12 signifies that the payoffs have been best optimised by the decision theoretic system in place and that the optimum value was reached in case of the threshold being kept at 9. This result and all the others before consistently demonstrate the efficiency of our decision theoretic model in filtering possible defectors and increasing cooperativeness by encouraging cooperation among the players.

7. CONCLUSION AND FUTUREWORK

In this paper we presented the model of a decision theoretic trustworthiness assessment system and plugged that in to our experimental framework to study its effectiveness in increasing co-operation among the transacting players in the game theoretic setting. The model based on the principles of the maximum expected utility besides being realistic was seen to be effective in cooperation building while preserving the average gain in each transaction. Game models of anonymous interactions in online business environments involve money as a part of transaction. Equating value of money with the price of goods- though justifiable- is not that convincing. If the value of money and price of the goods had same level of importance, then there would perhaps be an equilibrium condition and no exchange of goods would occur between the seller and buyer. In reality, money is more important to the seller and goods more important to buyer, and hence they transact. To represent utility of money in such cases, we have suggested using a utility based model like the one we presented in this paper. Further interesting models that can root down the desire to default can be formulated using these types of models. This forms a basis for future work in this area. The model developed can be scaled to find its application in scenarios like that of comparing products in online stores. A decision theoretic model like this could help users to choose a particular product on the basis of its maximum expected utility. MEUs consider both the value of goods and the trustworthiness of other party, thus making it easier for the user to make a decision. A limitation of the current model is that it can only handle goods with same value. In real world scenario, products being compared could have a range of price value. The existing model could be enhanced to compute MEUs for multi-valued products that lie in the same range of price. Another possible application area for this model is in recommender systems for social networks. In such cases, the value of goods referred here could be the rating of an object, while the cooperation probability could be an interpretation of the reputation of the recommender. A decision can thus be made on the basis of reliability of the object and that of the recommender.

REFERENCES

Aberer, K. and D. Zoran 2006. The Complex Facets of Reputation and Trust. 9th Fuzzy Days, International conference on computational intelligence Fuzzy Logic Neural Networks and Evolutionary Algorithms, Dortmund, Germany.

Aljazzaf, Z. M.and M. Perry 2010. Trust in Web Services. Services (SERVICES-1), 2010 6th World Congress on.

Avinash, L. V. and Ashwin S. 2008. A decision theoretic framework for analyzing binary hash-based content identification systems. Proceedings of the 8th ACM workshop on Digital rights management. Alexandria, Virginia, USA, ACM.

Axelrod, R. and Ed. 1984. The Evolution of Cooperation, Basic Books, New York.

Axelrod, R.and Ed. 1987. The Evolution of Strategies in the Iterated Prisoner's Dilemma. Genetic Algorithms and Simulated Annealing. San Francisco, CA, USA, Morgan Kaufmann Publishers Inc.

Ba, S. and A. B. Whintson 2003. "Building Trust in online auction markets through an economic incentive mechanism." Decision Support Systems 35: 273-286.

Berger, J. O. 1985. Statistical Decision Theory and Bayesian Analysis. New York, Springer-Verlag New York Inc.

Bianconi, M. 2003. Financial economics, risk and information: an introduction to methods and models Singapore, World Scientific Publishing Co.Pte Ltd.

Bista, S. K.and K. P. Dahal et al. 2008. Unravelling the Evolution of Defectors in online Business Games. 2nd International Conference on Software Knowledge and Information Management (SKIMA) Kathmandu, SKIMA.

Blythe, J. 1999. "Decision Theoretic Planning." AI magazine 20(2): 37-54.

Chaudhuri, A. and B. Sopher, et al. 2002. "Cooperation in Social Dilemmas, trust and reciprocity." Journal of Economic Psychology 23: 231-249.

Chong, S. Y. and X. Yao 2007. "Multiple choices and Reputation in Multiagent Interactions." IEEE Transactions on Evolutionary Computation 11(6): 689-710.

Cohen, W. W. and R. E. Schapire, et al. 1998. "Learning to order things." Advances in Neural Information Processing, Morgan Kaufmann 10.

Daniel Olmedilla , R. L.,Axel Polleres and Holger Lausen 2004. Trust Negotiation for Semantic Web Services. Semantic Web Services and Web Process Composition-SWSWPC. A. P. S. Jorge Cardoso. San Diego, CA, USA.

Dragoni, N. 2010. A Survey on Trust-Based Web Service Provision Approaches. Dependability (DEPEND), 2010 Third International Conference on.

Errity, A. 2003. Evolving Strategies for Prisoner's dilemma. Dublin, Dublin City University.

Frean, M. R. and E. R. Abraham 2001. "A Voter Model of the Spatial Prisoner's Dilemma " IEEE Transactions on Evolutionary Computation 5(2): 117-121.

Friedman, M. and L. Savage 1948. "The Utility Analysis of Choices Involving Risk." The Journal of Political Economy 56(4): 279-304.

Goldbeck, J.2002. Evolving Strategies for the Prisoner's Dilemma. Advances in Intelligent Systems, Fuzzy Systems, and Evolutionary Computation: 299-306.

Goldberg, D. E. 1989. Genetic Algorithms in search, Optimization and Machine Learning, Pearson Education Inc.

Hansson, S. O.1994. Decision theory a brief Introduction. Risk Analysis Course Material. Stockholm, Department ot Philosophy and the History of Technology, Royal Institute of Technology(KTH).

Holland, J.1975. Adaptation in Natural and Artificial Systems. Ann Arbor, University of Michigan Press.

Horvitz, E. J.and J. S. Breese, et al.1988. "Decision Theory in Expert Systems and Artificial Intelligence." International Journal of Approximate Reasoning 2(Special Issue on uncertainity in AI): 247-302.

Howley, E.and C. o. Riordan2007. "The effects of viscosity in choice and refusal IPD environments " Artificial Intelligence Review Springer Netherlands 26(1-2): 103-114.

Janssen, M.2006. "Evolution of Cooperation when feedback to reputation scores is voluntary." Journal of Artificial Scocieties and Social Simulation 9(1): 17.

Jordaan, I. J.2005. The Concept of Utility. Decisions under uncertainity: probabilistic analysis for engineering decisions. Cambridge, Cambridge Univer: 162-219.

JURCA, R.AND B. FLATINGS2004. An Incentive Compatible Reputation Mechanism. IEEE Conference on Electronic Commerce, CA, USA, IEEE Computer Society.

KAHNEMAN, D.AND A. TVERSKY1979. “Prospect Theory: An Analysis of Decision Under Risk.” Econometrica 47(2): 263-292.

KAMVAR, S. D.AND M. T. SCHOLSSER, ET AL.2003. The Eigen Trust Algorithm for Reputation Management in P2P Networks. ACM WWW2003, Budapest, Hungary, ACM.

KANEKO, M.AND M. H. WOODERS2004. Handbook of Utility Theory Volume 2. Dordrecht, The Netherlands, Kluwer Academic Publishers.

KRITZMAN, M.1998. Risk and utility:Basics. Investment Management. P. L. Bernstein and A. Damodaran. New York, John Wiley and Sons: 27-99.

KUHN, S.2007. “Prisoner’s Dilemma.” The Stanford Encyclopedia of Philosophy Winter 2007 Edition, from http://plato.stanford.edu/archives/win2007/entries/prisoner-dilemma/.

LI, J.-W.2004. “What Determines a Game to be Cooperative or Non-Cooperative?” SSRN eLibrary.

MALIK, Z.AND A. BOUGUETTAYA2009. “Reputation Bootstrapping for Trust Establishment among Web Services.” Internet Computing, IEEE 13(1): 40-47.

MASUDA, N.AND K. AIHARA2003. “Spatial prisoner’s dilemma optimally played in small-world networks.” Elsevier Physics Letters A(313): 55-61.

NEPAL, S.AND Z. MALIK, ET AL.2011. “Reputation Management for Composite Services in Service-Oriented Systems.” International journal of web services research 8(2): 29-52.

PARSONS, S.AND M. WOOLDRIDGE2002. “Game Theory and Decision Theory in Multi-Agent Systems.” Autonomous Agents and Multi-Agent Systems 5(3): 243-254.

POLI, R.AND W. B. LANGDON1997. Genetic Programming with one-Point Crossover. Second On-Line World Conference on Soft Computing in Engineering Design and Manufacturing, Springer-Verlag London.

RIEGELSBERGER, J.AND M. A. SASSE, ET AL.2005. “The mechanics of trust: A framework for research and design.” International Journal of Human-Computer Studies 62: 381-422.

ROSS, D.2009. “Game Theory.” The Stanford Encyclopedia of Philosophy (Fall 2009 Edition). Retrieved 2nd October 2009, 2009, from http://plato.stanford.edu/archives/fall2009/entries/game-theory/.

RUSSELL, S. J.AND P. NORVIG1998. Artificial intelligence : a modern approach. Upper Saddle River, N.J., Prentice Hall.

RUSSELL, S. J.AND P. NORVIG2010. Artificial intelligence : a modern approach. Upper Saddle River, N.J., Prentice Hall.

and (2006).

WEISSTEIN, E. W.2009. “Fair Game.” MathWorld–A Wolfram Web Resource. Retrieved July 7 2009, 2009, from http://mathworld.wolfram.com/FairGame.html

WEST, A. G.AND J. CHANG, ET AL.2011. “Trust in collaborative web applications.” Future Generation Computer Systems(0).

XIONG, L.AND L. LIU 2004. “PeerTrust: Supporting Reputation Based Trust for Peer-to-Peer electronic Communities.” IEEE Transactions on Knowledge and Data Engineering 16(7): 843-857.

YE DIANA, W.AND F. GUISSEPPI2006. “A decision-theoretic approach to the evaluation of information retrieval systems.” Inf. Process. Manage. 42(4): 863-874.

YU, B.AND M. P. SINGH2000. A Social Mechanism of Reputation Management in Electronic Communitoes. 4th International Workshop on Cooperative Information Agents(CIA) Berlin, Springer-Verlag.

YU, B.AND M. P. SINGH2002. Search in Referral Networks. AAMAS Workshop on Regulated Agent-Based Social Systems: Theories and Applications, 2002, Bologna, Italy.

ZHOU, R.AND K. HWANG2007. “PowerTrust: A Robust and Scalable Reputation System for Trusted Peer-to-Peer Computing.” IEEE Transactions on Parallel and Distributed Systems 18(4): 460-473.

ZOETER, O.AND M. TAYLOR, ET AL.2008. A Decision Theoretic Framework for Ranking using Implicit Feedback. SIGIR 2008 Workshop on Learning to Rank for Information Retrieval. Singapore.

Dr. Sanat Bista is a Research Fellow in ICT with the CSIRO, Australia. His research interest lies in the area of Trust, Privacy and Security in settings such as Social Networks, E-Commerce, Mobile Ad-Hoc Networks and the Cloud. His PhD research at the University of Bradford-UK involved modelling a game theoretic framework to asses the trustworthiness of interacting players in e-commerce and mobile ad-hoc network-like settings.
Prior to joining CSIRO, he was an Assistant Professor with the Department of Computer Science and Engineering at Kathmandu University, Nepal.

Dr. Keshav Dahal is a Reader (Associate Professor) in computational intelligence in the AI Research Centre at University of Bradford, UK. Prior to joining Bradford in 2002 he was a research fellow with the University of Strathclyde in Scotland, where he also obtained his PhD and MSc degrees. His research interests lie in the areas of decision support technologies, AI applications, Evolutionary computation. Dr Dahal has extensively published in peer-reviewed journals and conferences proceedings; and edited four books published by Springer. He is a member of the Institute of Engineering Technology (IET) and IEEE. He has sat on the organising and programme committees of many international conferences. He has been a visiting academic to Chiang Mai University (Thailand), Kantipur Engineering College (Nepal) and Universite Lumiere Lyon 2 (France).

Peter Cowling is Professor of Computer Science and Associate Dean (Research and Knowledge Transfer) at the University of Bradford (UK), where he leads the Artificial Intelligence research centre. He received the MA and DPhil degrees in Mathematical Sciences from Corpus Christi College of the University of Oxford, UK, in 1989 and 1997, respectively. His work centres on computerised decision-making in games, scheduling and resource-constrained optimisation. He has worked with a wide range of industrial partners, is director of 2 research spin-out companies, and has published over 80 scientific papers in high-quality journals and conferences. Prof. Cowling is a founding Associate Editor of the IEEE Transactions on Computational Intelligence and AI in Games.